

# Multidimensional cellular automata and generalization of Fekete's lemma

Silvio Capobianco\*

## Abstract

Fekete's lemma is a well known combinatorial result on number sequences: we extend it to functions defined on  $d$ -tuples of integers. As an application of the new variant, we show that nonsurjective  $d$ -dimensional cellular automata are characterized by loss of arbitrarily much information on finite supports, at a growth rate greater than that of the support's boundary determined by the automaton's neighbourhood index.

Keywords: subadditive function, product ordering, cellular automaton  
*Mathematics Subject Classification 2000:* 00A05; 37B15; 68Q80.

## 1 Introduction

Let  $f : \{1, 2, \dots\} \rightarrow [0, +\infty)$ . **Fekete's lemma** [5, 8, 10] states that, if  $f(n+k) \leq f(n) + f(k)$  for every  $n$  and  $k$ , then

$$\lim_{n \rightarrow \infty} \frac{f(n)}{n} = \inf_{n \geq 1} \frac{f(n)}{n}. \quad (1)$$

The consequences of this simple statement are many and deep, such as the definition of *topological entropy* for dynamical systems [1, 7, 8] and Arratia's bound on the number of permutations avoiding a given pattern [3].

More recently, in a joint work with Tommaso Toffoli and Patrizia Menestrà [9], we have made use of (1) to prove a result on unidimensional, non-surjective *cellular automata* (CA). CA are presentations of global dynamics in local terms: each global state is a  $d$ -dimensional *configuration*, and the global evolution rule changes the state locally at a site by considering only the states of *neighbouring* sites. Nonsurjective CA are characterized by the existence of a configuration *on a finite region of the space* that has no predecessor according to the evolution rule; this can be restated from another point of view, by saying that nonsurjective CA *lose information within finite range*. Fekete's lemma then told us that such CA must lose, on finite regions large enough, an amount of information essentially proportional to the size of the support itself—thus, at least

---

\*School of Computer Science, Reykjavík University; email: [silvio@ru.is](mailto:silvio@ru.is)  
 The author was partly supported by the project “The Equational Logic of Parallel Processes” (nr. 060013021) of The Icelandic Research Fund.

the size of the support's *boundary* determined by the neighborhood index; since loss of information gives rise to *extra channel capacity*, our team has devised a general algorithm to translate from a presentation using an  $n$ -inputs, 1-output local map (*i.e.*, CA) to one employing  $n$ -inputs,  $n$ -outputs events, characteristic of a different class of presentations, specifically, that of *lattice gases* (LG).

In this paper, we state and prove a multivariate version of Fekete's lemma. The motivation for this, is to provide a support to the conjecture that the translation algorithm in [9] could be extended to arbitrary dimension. To prove our generalization, we rearrange a proof of (1) so that it works on sequences of integer  $d$ -tuples, after a suitable ordering on these is defined. After that, we use the more general result at our hands to show that the same phenomenon that allows rewriting CA as LG in dimension 1, actually occurs in arbitrary finite dimension. Incidentally, we get a criterion for CA surjectivity.

## 2 Fekete's lemma, multivariate

Let  $\mathbb{Z}_+ = \mathbb{Z} \cap (0, +\infty)$ . Consider the **product ordering** on  $\mathbb{Z}_+^d$  defined by  $x \leq_\pi y$  iff  $x_i \leq y_i$  for every  $i \in \{1, \dots, d\}$ : this is the kind of ordering used, *e.g.*, in *linear programming*, by writing  $Ax \leq b$  to indicate a set of constraints  $a_{1,1}x_1 + \dots + a_{1,n}x_n \leq b_1, \dots, a_{m,1}x_1 + \dots + a_{m,n}x_n \leq b_m$ ; it is also the *finest* ordering that makes the projections *monotonic*. Observe that  $\mathcal{Z}^d = (\mathbb{Z}_+^d, \leq_\pi)$  is a **directed set**, *i.e.*, for any two  $x, y \in \mathbb{Z}_+$  there exists  $z \in \mathbb{Z}_+$  such that both  $x \leq_\pi z$  and  $y \leq_\pi z$ . If  $\mathcal{X} = (X, \leq)$  is a directed set and  $f : X \rightarrow \mathbb{R}$ , the **lower** and **upper limit** of  $f$  in  $\mathcal{X}$  are defined as usual, *i.e.*,

$$\liminf_{x \in \mathcal{X}} f(x) = \sup_{x \in \mathcal{X}} \inf_{y \geq x} f(y) \quad \text{and} \quad \limsup_{x \in \mathcal{X}} f(x) = \inf_{x \in \mathcal{X}} \sup_{y \geq x} f(y) ;$$

moreover,  $f$  has **limit**  $L \in \mathbb{R}$  in  $\mathcal{X}$ , written  $\lim_{x \in \mathcal{X}} f(x) = L$ , if for every  $\varepsilon > 0$  there exists  $x_\varepsilon \in X$  such that  $|f(x) - L| < \varepsilon$  for every  $x \geq x_\varepsilon$ . For example, if  $r_1, \dots, r_d \in \mathbb{N} = \mathbb{Z}_+ \cup \{0\}$  are fixed, then

$$\lim_{(x_1, \dots, x_d) \in \mathcal{Z}^d} \frac{(x_1 + r_1) \cdots (x_d + r_d)}{x_1 \cdots x_d} = 1 . \quad (2)$$

It follows from the definitions that  $\liminf_{x \in \mathcal{X}} f(x) \leq \limsup_{x \in \mathcal{X}} f(x)$ , and that  $\lim_{x \in \mathcal{X}} f(x) = L$  iff  $\liminf_{x \in \mathcal{X}} f(x) = \limsup_{x \in \mathcal{X}} f(x) = L$ .

**Theorem 1** *Let  $f : \mathbb{Z}_+^d \rightarrow [0, +\infty)$  satisfy*

$$f(x_1, \dots, x_j + y_j, \dots, x_d) \leq f(x_1, \dots, x_j, \dots, x_d) + f(x_1, \dots, y_j, \dots, x_d) \quad (3)$$

*for every  $x_1, \dots, x_n, y_j \in \mathbb{Z}_+$ ,  $j \in \{1, \dots, d\}$ . Then*

$$\lim_{(x_1, \dots, x_d) \in \mathcal{Z}^d} \frac{f(x_1, \dots, x_d)}{x_1 \cdots x_d} \quad (4)$$

*exists, and equals*

$$\inf_{x_1, \dots, x_d \in \mathbb{Z}_+} \frac{f(x_1, \dots, x_d)}{x_1 \cdots x_d} .$$

*Proof.* Because of (3), for every  $j \in \{1, \dots, d\}$ ,  $x_1, \dots, x_d \in \mathbb{Z}_+$ , if  $x_j = qt + r$  with  $q \in \mathbb{N}$  and  $r \in \mathbb{Z}_+$ , then

$$f(x_1, \dots, x_j, \dots, x_d) \leq qf(x_1, \dots, t, \dots, x_d) + f(x_1, \dots, r, \dots, x_d). \quad (5)$$

Fix  $t_1, \dots, t_d \in \mathbb{Z}_+$ . For each  $(x_1, \dots, x_d) \in \mathbb{Z}_+^d$ ,  $d$  pairs  $(q_j, r_j) \in \mathbb{N} \times \mathbb{Z}_+$  are uniquely determined by  $x_j = q_j t_j + r_j$  and  $1 \leq r_j \leq t_j$ . By repeatedly applying (5) to all of the  $x_j$ 's we find

$$\begin{aligned} f(x_1, \dots, x_d) &\leq q_1 \cdots q_d f(t_1, \dots, t_d) \\ &\quad + q_1 \cdots q_{d-1} f(t_1, \dots, t_{d-1}, r_d) + \dots \\ &\quad + q_1 \cdots q_{d-2} f(t_1, \dots, t_{d-2}, r_{d-1}, r_d) + \dots \\ &\quad + \dots \end{aligned} \quad (6)$$

where, in the next  $k$ 'th line,  $k \geq 1$ , each occurrence of  $f$  has  $k$  arguments chosen from the  $r$ 's and  $d - k$  chosen from the  $t$ 's, and is multiplied precisely by the  $q$ 's corresponding to the  $t$ 's; moreover, all these occurrences are bounded from above by the constant  $M = t_1 \cdots t_d \cdot f(1, \dots, 1)$ . Divide both sides of (6) by  $x_1 \cdots x_d$ : since  $\lim_{x_j \rightarrow \infty} q_j / x_j = 1/t_j$ , if all the  $x_j$ 's are large enough, then the first summand of right-hand side becomes very close to  $f(t_1, \dots, t_d)/t_1 \cdots t_d$ , and the other ones become very small; that is, for every  $\varepsilon > 0$ , there exists  $(x_1, \dots, x_d) \in X$  such that, for every  $(y_1, \dots, y_d) \geq_\pi (x_1, \dots, x_d)$ ,

$$\frac{f(y_1, \dots, y_d)}{y_1 \cdots y_d} < \frac{f(t_1, \dots, t_d)}{t_1 \cdots t_d} + \varepsilon.$$

From this follows

$$\limsup_{(x_1, \dots, x_d) \in \mathcal{Z}^d} \frac{f(x_1, \dots, x_d)}{x_1 \cdots x_d} \leq \frac{f(t_1, \dots, t_d)}{t_1 \cdots t_d};$$

this is true whatever the  $t_j$ 's are, hence

$$\limsup_{(x_1, \dots, x_d) \in \mathcal{Z}^d} \frac{f(x_1, \dots, x_d)}{x_1 \cdots x_d} \leq \inf_{t_1, \dots, t_d \in \mathbb{Z}_+} \frac{f(t_1, \dots, t_d)}{t_1 \cdots t_d}.$$

The thesis then follows from the inequality

$$\inf_{t_1, \dots, t_d \in \mathbb{Z}_+} \frac{f(t_1, \dots, t_d)}{t_1 \cdots t_d} \leq \liminf_{(x_1, \dots, x_d) \in \mathcal{Z}^d} \frac{f(x_1, \dots, x_d)}{x_1 \cdots x_d}.$$

□

### 3 An application to cellular automata

A **cellular automaton** (briefly, CA) is a quadruple  $\mathcal{A} = \langle d, Q, \mathcal{N}, f \rangle$  where the **dimension**  $d > 0$  is an integer, the **set of states**  $Q$  is finite and has at least two distinct elements, the **neighbourhood index**  $\mathcal{N} = \{\nu_1, \dots, \nu_n\}$  is a

finite subset of  $\mathbb{Z}^d$ , and the **local evolution function**  $f$  maps  $Q^n$  into  $Q$ . A **global evolution function**  $F$  is induced by  $f$  of the space  $Q^{\mathbb{Z}^d}$  of  $d$ -dimensional configurations by

$$(F(c))(x) = f(c(x + \nu_1), \dots, c(x + \nu_n)) . \quad (7)$$

$\mathcal{A}$  is said to be surjective if  $F$  is. For example, if  $d = 1$ ,  $Q = \{0, 1\}$ ,  $\mathcal{N} = \{+1\}$ ,  $f(x) = x$ , then  $\langle d, Q, \mathcal{N}, f \rangle$  is the *shift cellular automaton* and  $(F(c))(x) = c(x + 1)$  is the *shift map*, which is surjective; on the other hand, for same  $d$  and  $Q$ ,  $\mathcal{N} = \{0, +1\}$ , and  $f(a, b) = a \cdot b$ , we get a nonsurjective CA, because if  $\bar{c}(x)$  is 0 for  $x = 0$  and 1 otherwise, then  $F(c) \neq \bar{c}$  for any  $c$ .

For every finite  $E \subseteq \mathbb{Z}^d$ , calling  $E + \mathcal{N} = \{x + \nu \mid x \in E, \nu \in \mathcal{N}\}$ , a function  $F_E : Q^{E+\mathcal{N}} \rightarrow Q^E$  is induced by  $f$ , again by applying (7). Observe that the number  $|F_E(Q^{E+\mathcal{N}})|$  of **patterns** over  $E$  obtainable by applying (7) does not depend on the *displacement* of  $E$  along  $\mathbb{Z}^d$ , i.e., if  $x + E = \{x + y \mid y \in E\}$ , then  $|F_{x+E}(Q^{x+E+\mathcal{N}})| = |F_E(Q^{E+\mathcal{N}})|$ . It is well known (cf. [4]) that  $\mathcal{A}$  is surjective iff  $F_E$  is surjective for every  $E$  which is a **right  $d$ -polytope**, i.e., a subset of  $\mathbb{Z}^d$  of the form  $\prod_{i=1}^d \{k_i, \dots, k_i + s_i - 1\}$ ,  $s_1, \dots, s_d \in \mathbb{Z}_+$  being the **sides**. (Here, “right” has the same meaning as in “right triangle”.) Put  $E(x_1, \dots, x_d) = \prod_{i=1}^d \{0, \dots, x_i - 1\}$  : if  $\mathcal{N}$  is contained in a right  $d$ -polytope of sides  $r_1, \dots, r_d$ , then  $E(x_1, \dots, x_d) + \mathcal{N}$  is contained in a right  $d$ -polytope of sides  $x_1 + r_1, \dots, x_d + r_d$ , which is the disjoint union of  $E(x_1, \dots, x_d)$  and a **boundary**.

Let  $\mathcal{A} = \langle d, Q, \mathcal{N}, f \rangle$  be a CA. If  $\mathcal{A}$  is nonsurjective, then there must exist a support of suitable size where not every possible pattern is reachable, i.e., a part of the information is lost. In the 1D case [9], such lost information is proved to ultimately be as much as the boundary can transport; which allowed devising a CA-to-LG conversion algorithm. If the technique employed there is to be extended to higher dimension, then we must determine whether such large a loss can still be achieved.

Call **output size** of  $f$  on a right  $d$ -polytope of sides  $x_1, \dots, x_d$  the quantity

$$\text{Out}_f(x_1, \dots, x_d) = \left| F_{E(x_1, \dots, x_d)} \left( Q^{E(x_1, \dots, x_d) + \mathcal{N}} \right) \right| .$$

Then  $\mathcal{A}$  is surjective iff  $\text{Out}_f(x_1, \dots, x_d) = |Q|^{x_1 \cdots x_d}$  for every  $x_1, \dots, x_d \in \mathbb{Z}_+$ . By switching to a logarithmic measure unit, we can associate to  $\mathcal{A}$  a **loss of information** on a right  $d$ -polytope of sides  $x_1, \dots, x_d$  defined as

$$\Lambda_{\mathcal{A}}(x_1, \dots, x_d) = x_1 \cdots x_d - \log_{|Q|} \text{Out}_f(x_1, \dots, x_d) . \quad (8)$$

Observe how such loss is measured in *qits* (with  $q = |Q|$ ), a *qit* being the amount of information carried by a  $q$ -states device;  $n$  *qits* correspond to  $n \log_2 q$  bits.

**Theorem 2** *Let  $\mathcal{A} = \langle d, Q, \mathcal{N}, f \rangle$  be a CA. Define  $\Lambda_{\mathcal{A}}$  as by (8). Then*

1. *either  $\mathcal{A}$  is surjective and  $\Lambda_{\mathcal{A}}$  is identically zero,*

2. or  $\mathcal{A}$  is nonsurjective and for every  $K \geq 0$ ,  $r_1, \dots, r_d \in \mathbb{N}$ , there exist  $t_1, \dots, t_d \in \mathbb{Z}_+$  such that, for every  $(x_1, \dots, x_d) \geq_\pi (t_1, \dots, t_d)$ ,

$$\Lambda_{\mathcal{A}}(x_1, \dots, x_d) \geq (x_1 + r_1) \cdots (x_d + r_d) - x_1 \cdots x_d + K.$$

In particular, if  $\Lambda_{\mathcal{A}}$  is bounded, then  $\mathcal{A}$  is surjective.

*Proof.* Put  $q = |Q|$ . Since a pattern over  $E(x_1, \dots, x_j + y_j, \dots, x_d)$  can always be seen as the *joining* of a pattern over  $E(x_1, \dots, x_j, \dots, x_d)$  and another one over  $E(x_1, \dots, y_j, \dots, x_d)$ , there cannot be more patterns obtainable over the former than pairs of patterns obtainable over the latter, *i.e.*,

$$\text{Out}_f(x_1, \dots, x_j + y_j, \dots, x_d) \leq \text{Out}_f(x_1, \dots, x_j, \dots, x_d) \cdot \text{Out}_f(x_1, \dots, y_j, \dots, x_d)$$

whatever  $x_1, \dots, x_n, y_j \in \mathbb{Z}_+$ ,  $j \in \{1, \dots, d\}$  are; consequently,  $\log_q \text{Out}_f$  is subadditive in each of its arguments (and nonnegative). Let

$$\lambda_f = \lim_{(x_1, \dots, x_d) \in \mathbb{Z}^d} \frac{\log_q \text{Out}_f(x_1, \dots, x_d)}{x_1 \cdots x_d}, \quad (9)$$

whose existence and value are given by Theorem 1; observe that  $\lambda_f \leq 1$ , and  $\mathcal{A}$  is surjective iff  $\lambda_f = 1$ . Suppose  $\mathcal{A}$  is nonsurjective. Let  $\delta \in (\lambda_f, 1)$ . Choose  $t_1, \dots, t_d \in \mathbb{Z}_+$  so that, for every  $(x_1, \dots, x_d) \geq_\pi (t_1, \dots, t_d)$ , both

$$\frac{\log_q \text{Out}_f(x_1, \dots, x_d)}{x_1 \cdots x_d} \leq \delta \quad (10)$$

and

$$\frac{(x_1 + r_1) \cdots (x_d + r_d) - x_1 \cdots x_d + K}{x_1 \cdots x_d} \leq 1 - \delta \quad (11)$$

are satisfied, the latter following from (2). Then, for such  $x_1, \dots, x_d$ ,

$$\begin{aligned} x_1 \cdots x_d - \log_q \text{Out}_f(x_1, \dots, x_d) &\geq (x_1 \cdots x_d)(1 - \delta) \\ &\geq (x_1 + r_1) \cdots (x_d + r_d) - x_1 \cdots x_d + K. \end{aligned}$$

□

From Theorem 2 follows that, for  $(x_1, \dots, x_d)$  satisfying both (10) and (11), *the loss of information is at least the size of the boundary*: this is precisely the fact used in [9], and supports the conjecture that a similar construction can be carried out in dimension  $d > 1$ . On the other hand—and perhaps, unfortunately—since surjectivity of  $d$ -dimensional CA is only decidable when  $d = 1$  [2, 6], no algorithm exists to determine, given an arbitrary multidimensional CA, that its loss of information (8) is bounded.

## Acknowledgements

The proof of Theorem 1 is an adaptation of an argument shown to us by Tullio Ceccherini–Silberstein. We also thank Tommaso Toffoli, Patrizia Mentrasti, Luca Aceto, Anna Ingólfssdóttir, Anders Claesson, and Magnús Már Halldórsson for the many helpful suggestions and encouragements.

## References

- [1] R. L. Adler, A. G. Konheim, M. H. McAndrew. Topological entropy. *Trans. Am. Math. Soc.* **114** (1965) 309–319.
- [2] S. Amoroso, Y. N. Patt. Decision procedures for surjectivity and injectivity of parallel maps for tessellation structures. *J. Comput. System Sci.* **6** (1972) 448–464.
- [3] R. Arratia. On the Stanley-Wilf Conjecture for the Number of Permutations Avoiding a Given Pattern. *Elec. J. Comb.* **6** (1999) N1.
- [4] C. Calude. *Information and Randomness. An Algorithmic Perspective*. 2nd ed., Springer 2002.
- [5] M. Fekete. Über die Verteilung der Wurzeln bei gewisser algebraischen Gleichungen mit ganzzahligen Koeffizienten. *Math. Zeitschr.* **17** (1923) 228–249.
- [6] J. Kari. Reversibility of 2D cellular automata is undecidable. *Physica D* **45** (1990), 379–385.
- [7] P. Koiran. The Topological Entropy of Iterated Piecewise Affine Maps is Uncomputable. *Disc. Math. Theor. Comp. Sci.* **4** (2001) 351–356.
- [8] D. Lind, B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press 1995.
- [9] T. Toffoli, S. Capobianco, P. Mentrasti. When—and how—can a cellular automaton be rewritten as a lattice gas? [arxiv.org/abs/0709.1173](https://arxiv.org/abs/0709.1173), submitted to *Theor. Comp. Sys.*
- [10] J.L. van Lint, R.M. Wilson. *A Course in Combinatorics*. Cambridge University Press 1992.